DOCUMENT RESUME

ED 236 224 TM 830 745

AUTHOR Wilson, Morris D.
TITLE The Assessment of Composition Skills. Pre/Post Results: Group I.
INSTITUTION Des Moines Public Schools, IA. Dept. of Evaluation and Research.
PUB DATE 22 Jul 83
NOTE 41p.; Faint print in some pages of appendices.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Achievement Gains; Educational Assessment; Holistic Approach; *Pretests Posttests; Program Effectiveness; Scoring; Secondary Education; *Writing Evaluation; *Writing Improvement; Writing Skills
IDENTIFIERS Des Moines Public Schools IA

ABSTRACT

The initial stimulus for measuring the composition skills of students in the Des Moines (IA) Independent Community School District came in 1980-81 when the school board adopted a goal for the year which stated: "The district should have school programs that improve pupils' competency in written expression." The intent of the original writing assessment evaluation design was to provide pre/post assessments for two groups of students in order to determine if satisfactory growth in composition skills was being realized. The instrument, a copy of which is in Appendix A, was administered to 2,198 ninth-grade students in September 1981, and to the same students in May 1983. The 1982-83 freshman class has also been pre-tested, and will be post-tested in May 1984. Each student's essay was scored holistically and in terms of six specific composition traits (holistic, development, organization, language choice, sentence structure, and mechanics). The average scores across all six scales showed moderate overall gains for the first group of students, but further analysis revealed that around 40 percent of the individual students failed to show any progress on five of the six score scales. (BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED236224

THE ASSESSMENT OF COMPOSITION SKILLS

Pre/Post Results: Group I

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Morris D. Wilson, Ph.D.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. D. Wilson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Department of Evaluation and Research
Des Moines Independent Community School District
1800 Grand Avenue
Des Moines, IA 50307

TM830 745

Approved: Morris D. Wilson

Morris D. Wilson, Ph.D.
Director of Evaluation and Research

July 22, 1983

READING GUIDE

Because of the nature of the type of assessment/testing which was done and is reported in this document, no "abstract" can be provided. However, to assist the reader, attention to the following is recommended:

1. Background and introduction ..... Page 1

2. Costs ..... Page 5

3. The instrument itself ..... Appendix A, Page 19

4. Methods of scoring ..... Page 3

5. The narrative criteria (rubrics) for scoring ..... Appendix B, Pages 24-26

6. Number of students taking the test ..... Page 7

7. The prototype papers and their importance in scoring ..... Pages 5-6 and Appendix C, Page 28+

8. Results ..... Pages 8-14

9. Comments ..... Pages 15-17

INTRODUCTION

The initial stimulus for measuring the composition skills of students in the Des Moines Independent Community School District came in 1980-81 when the school board adopted a goal for the year which stated: "The district should have school programs that improve pupils' competency in written expression." The Department of Evaluation and Research responded to the goal by developing an instrument to assess the writing proficiency of district students.

The intent of the original writing assessment evaluation design was to provide pre/post assessments for two groups of students in order to determine if satisfactory growth in composition skills was being realized.

The instrument, a copy of which is placed in Appendix A, was first administered during the week of September 14-18, 1981. A total of 2,198 ninth grade students (approximately 91 percent of the total ninth grade enrollment) took the writing assessment at that time. This class, (the class of 1985) now tenth graders, was tested again in May, 1983, with the same instrument. This report details the results of pre/post testing obtained from the two administrations of the instrument.

The second group of students, the 1982-83 freshman class, has been pretested at the beginning of their ninth grade year and will be post-tested at the end of their tenth grade year in May, 1984. That administration will complete the scheduled plan for writing assessment in the district. If an extension of the writing assessment program is desired by the district, that decision should be made early during the 1983-84 school year so appropriate planning (rubrics, instruments and adequate instrument validation) may be accomplished.

METHODOLOGY

All the writing assessments were administered by senior high classroom teachers in their regularly scheduled English classes. It was determined that this arrangement was more conducive to obtaining a representative sample of a students' writing ability, since it was a natural setting for the students and also familiarized the classroom teachers through direct exposure with what the writing sample was attempting to determine. Students were allowed 45 minutes for planning, writing and recopying their essays. Other instructions may be found inside the Writing Assessment Booklet (Appendix A).

Scoring an actual writing sample is much more complex than scoring a standard multiple choice test. Standard testing instruments usually have a single correct answer for each test item; however, in the case of a writing sample, while the writing task and criteria for scoring can be standardized, there is no single 'correct answer.' Consequently, in order to facilitate the standardization of scoring (called "inter-rater reliability"), four components were necessary:

1. Criteria (rubrics) had to be specified and defined.
2. Prototype papers for each score on each criterion had to be identified.
3. The readers had to be trained to recognize the similarity between the prototype paper scores and the papers they were grading to a high degree of reliability.
4. Finally, in those instances where the readers disagreed by more than one score point the Chief Reader and Assistant Chief Reader needed to read the paper and assign a score.

The initial delineation of criteria was accomplished by a steering committee of teachers and the Supervisor of Language Arts during the process of instrument development. Prior to the actual scoring sessions the committee specifically defined each criterion in narrative form. The criteria can be seen on the front of the instrument (Appendix A) while the narrative definition of each criterion (rubrics) is shown in Appendix B.

SCORING

The reader should observe that there were two major areas of scoring: (1) Holistic and (2) Composition Traits. The holistic scores were obtained from the overall impression the paper made upon the reader, while the composition traits were five specific traits identified by the steering committee which comprise the various elements to which the writer must attend if he/she is to succeed in developing a satisfactory composition (as defined by the steering committee).

Prototype papers were originally chosen by the steering committee. Each committee member scored several papers individually, using the holistic and the composition traits definitions. The members then agreed by consensus on which papers best modeled the various scores for each category.

All readers received two days of training in both holistic and composition trait scoring. The training was conducted by the Chief Reader (Doris Roettger) and the Assistant Chief Reader (Jack Owens), immediately prior to the beginning of writing assessment scoring. The readers all had teaching experience and/or experience in writing/composition. Thirteen readers were utilized for the first scoring session (November, 1981), while only nine were employed in May, 1983. The scorers were trained by first identifying and ranking typical papers of varying degrees of excellence, then moving into the practice of reading, scoring, and discussion, always striving toward higher degrees of similarity in scoring.

To insure continuing inter-rater reliability among the scorers, each paper was read by two randomly assigned readers. The scores of the first reader were marked on the front of the booklet, then masked so

that the second reader would not be aware of the first reader's scores. When both readers completed their scoring the papers were returned to the chief reader who unmasked the first reader's scores and compared them with those of the second reader. If the score for each reader was the same, or within one point (e.g., 5-4, 4-3, etc.), the scoring was considered final. If the scores were more than one point apart (e.g., 5-3, 4-2, etc.), the papers were then reread by the chief reader and assistant chief reader who determined the final score. According to the chief reader, the scoring had a high degree of inter-rater reliability and the rereading task was never overly-burdensome. As the half days of scoring progressed the reliability improved and both the chief reader and assistant chief reader were pleased with the quality of work exhibited by the readers.

The cost for scoring the writing assessments came to approximately $2,200 for the pretest and $1,475 for the posttest, or $3,675 for the complete pre/post series. This amounts to approximately $1 per student for the two administrations. Readers were paid $4.10 per hour. Teachers were not used to score the assessments because the cost for substitutes would more than double the scoring costs. However, if future assessments are made the district should consider if the scoring activity would be a beneficial inservice activity for English teachers and if some method can be devised to provide funds for that activity.

It was necessary not only to use prototype papers as a means of training scorers, but also as a means of interpreting scores. These papers provide a basis for comparison of scores so that one may assume that a paper (or average) is quite similar in quality to the various

prototypes. Some prototype papers for the holistic scores are found in Appendix C.

When the papers were read by two scorers and their scores were within one point of one another, the scores were then added to obtain the total score for that paper. The holistic scale thus consisted of 11 points (0-10). The prototype papers in Appendix C exhibit papers with a total holistic score of 10, 8, 6, 4, 2, and 0. This score is shown in the upper right corner of each paper. Using these as a basis of comparison, the reader may have a better grasp of the scores found in the tables accompanying this report.

Unless the reader closely examines the prototype papers in Appendix C, the figures contained in the report tables are meaningless. These are average scores of holistic and composition trait interpretations, NOT average scores of some firmly anchored numerical count (like number of items correct). The reader is reminded that the scores shown in the tables and on the prototype papers are the combined scores assigned by two readers for each paper.

One area of concern in the initial pretesting of the class of 1985 was the number of enrolled students who took the writing test. Appendix D (Table 5) shows that 91.4 percent of the students enrolled in the fall of 1981 took the test in September of that year. The concern expressed at that time was that unless students had both a pretest and posttest it would be impossible to measure the gains being exhibited by the group.

In May, 1983, this concern was compounded when only 85.9 percent of the students enrolled in tenth grade English classes during the spring semester, 1983, were administered the posttest. These figures are also shown in Appendix D (Table 6).

The individual student scores, pre, post, and gain, will be provided each building to assist in placement decisions and to make teachers aware of the accomplishments or needs of the students enrolled in their eleventh and twelfth grade English classes; however, students without complete pre and posttest results will not be included in the data. Senior high school building principals should make every effort to insure that the class of 1986 (1983-84 sophomores) are ALL included in the testing round to be conducted in May, 1984.

The problem of having all students take the composition assessment is even more acute when one examines the number of valid papers in more detail. Of the 2,198 papers scored in the fall 1981 assessment, 6 were not useable because of invalid identification numbers. This reduced the number of possible pre/post matches to 2,192. Of the 1,760 papers scored in the spring 1983 assessment, 95 had invalid (or no) student ID numbers, thus only 1,665 students had valid posttest papers which might have been matched with their pretests. Of the 1,665 posttest papers, 262 student ID numbers failed to match with ID numbers on the pretest, leaving a total of 1,403 students who had matching scores. Thus, only about 64 percent of the 2,198 students tested in fall 1981 had valid matching papers for pre/post comparison in May, 1983. Much of this discrepancy may be attributed to in/out migration, some was failure of teachers to write in the ID numbers on the posttest, and an additional number stemmed from the fact that not all students were administered either the posttest or pretest.

RESULTS

Table 1 displays the frequency and percent of districtwide scores occurring at each point on each of the six scales (holistic, development, organization, language choice, sentence structure, and mechanics), for both the pretest and the posttest, for those 1,403 students who had valid pre and posttest scores. Again, the same prototype papers were used to train the scorers and anchor the numerical scores for both administrations of this test. We can, therefore, assume that the same level of difficulty was applied in the scoring of each set of papers and that the pretest and posttest scores are comparable.

TABLE 3

District Total
Frequency and Percent of Scores Occurring at Each Point on Each of the Six Scales
Results from Fall, 1981 (N=1403) and Spring, 1983

Score	Year	Holistic		Development		Organization		Language Choice		Sentence Structure		Mechanics	
		#	%	#	%	#	%	#	%	#	%	#	%
0	1981	20	1.4	18	1.3	26	1.9	15	1.1	30	2.1	26	1.9
	1983	5	0.4	5	0.4	7	0.5	2	0.1	7	0.5	6	0.4
1	1981	15	1.1	21	1.5	24	1.7	13	0.9	50	3.6	38	2.7
	1983	12	0.9	9	0.6	10	0.7	5	0.4	12	0.9	19	1.4
2	1981	77	5.5	83	5.9	81	5.8	71	5.1	188	13.4	238	17.0
	1983	45	3.2	41	2.9	54	3.8	22	1.6	111	7.9	131	9.3
3	1981	103	7.3	113	8.1	135	9.6	109	7.8	263	18.7	288	20.5
	1983	69	4.9	61	4.3	81	5.8	68	4.8	173	12.3	193	13.8
4	1981	307	21.9	227	16.2	274	19.5	344	24.5	501	35.7	445	31.7
	1983	186	13.3	178	12.7	199	14.2	180	12.8	368	26.2	395	28.2
5	1981	215	15.3	233	16.6	231	16.5	283	20.2	229	16.3	234	16.7
	1983	199	14.2	205	14.6	191	13.6	213	15.2	317	22.6	326	23.2
6	1981	336	23.9	323	23.0	296	21.1	308	22.0	142	10.1	134	9.6
	1983	318	22.7	276	19.7	290	20.7	299	21.3	415	29.6	333	23.7
7	1981	170	12.1	203	14.5	180	12.8	146	10.4				
	1983	197	14.0	218	15.5	200	14.3	228	16.3				
8	1981	112	8.0	134	9.6	117	8.3	68	4.8				
	1983	188	13.4	211	15.0	176	12.5	166	11.8				
9	1981	35	2.5	33	2.4	33	2.4	33	2.4				
	1983	99	7.1	108	7.7	110	7.8	142	10.1				
10	1981	13	0.9	15	1.1	6	0.4	13	0.9				
	1983	85	6.1	91	6.5	85	6.1	78	5.6				

Upon examination of the results for the holistic, development, organization and language choice scales (scored from 0 to 10) one can observe that the percent of the scores falling in the 0 to 5 range declined, those scored at 6 remained about the same, and those scored from 7 to 10 increased. Thus, when this group of 1,403 students entered ninth grade, 23.5 percent of the students scored from 7 to 10 on the holistic scale. At the time they ended tenth grade this percentage had risen to 36.6. At the other end of the holistic scale, 52.5 percent scored from 0 to 5 on the pretest, while 36.9 percent were still at those levels on the posttest. From this viewpoint it would seem that the two years of instruction and maturity positively affected the holistic writing scores of the class of 1985 at all levels of prior achievement. The score improvements on the other 0 to 10 scales were similar to those for the holistic score, as shown in Table 2. The rubric describing each particular scale is shown in Appendix B and should be consulted as the reader attempts to interpret the scores shown in the tables.

TABLE 2

Gains in Percentage of Scores for Specific Scales
Class of 1985

	Holistic		Development		Organization		Language Development	
Scores	Pre	Post	Pre	Post	Pre	Post	Pre	Post
0 - 5	52.5	36.9	49.6	35.5	55.0	38.6	59.6	34.9
7 - 10	23.5	36.6	27.6	44.7	23.9	40.7	18.5	43.8

In the two scales scored from 0 to 6 (sentence structure and mechanics) similar gains were realized. In those areas the break point seemed to fall around four. In sentence structure, those scoring from

0 to 3 fell from 38.3 percent to 21.6 percent and those scoring from 5 to 6 rose from 26.4 percent to 52.2 percent. In mechanics, those scoring from 0 to 3 declined from 42.1 percent to 24.9 percent and those scoring from 5 to 6 rose from 26.3 percent to 46.9 percent.

From the group averages it is apparent that positive gains were made. However, whether these gains should have been even larger, or were as large as could be expected, are questions that always beg answers in a local assessment such as this. Without some kind of referrent group (norming group) against which to guage our progress we are typically forced to rely on informed judgment for a reply. In order to shed more light on the magnitude of gains shown by the individual members of the senior class of 1985 who had valid pre and post test ID's, Table 3 was produced to assist in the analysis.

Table 3 displays the individual student gains or losses from pretest to posttest for the 1,403 students with valid pre/post scores. The first column displays the amount of change (positive or negative). Then, for each of the six score scales, the number of students, percent of the 1,403 students, and cumulative percent of the 1,403 students is shown for that positive or negative change. For example, one can see that for the holistic scale no student's score dropped as much as 7 points, but one student dropped 6 points, which was 0.1 percent of the 1,403 students. Moving down the score change column, 4 students lost 5 points between pre and posttesting. These 4 students were 0.3 percent of the 1,403 students and the 5 students who lost 5 or more points were 0.4 percent of the 1,403 students.

Two hundred seventy-nine students had no change in their score between pre and posttesting. These 279 students represented 19.9

percent of the total 1,403 students. However, when we look at the cumulative percent column, we can see that 40.9 percent of the 1,403 students had either no gain, or a negative loss in score, between pre/posttest assessments. This is very disturbing. Apparently, the gains realized on the holistic scale and in Tables 1 and 2 come from only approximately 60 percent of the students, since 40.9 percent of the students had no gain, or regressed, between pretesting and posttesting.

TABLE 3

Individual Student Gains or Losses from Pretest to Posttest
Senior Class of 1985 - N=1403

Score Change + or -	Holistic			Development			Organization			Language Development			Sentence Structure			Mechanics		
	N	%	Cum%	N	%	Cum%	N	%	CUM%	N	%	CUM%	N	%	CUM%	N	%	CUM%
-7				1	0.1	0.1	1	0.1	0.1									
-6	1	0.1	0.1	1	0.1	0.1				1	0.1	0.1						
-5	4	0.3	0.4	2	0.1	0.3	3	0.2	0.3									
-4	11	0.8	1.1	13	0.9	1.2	16	1.1	1.4	5	0.4	0.4						
-3	24	1.7	2.9	38	2.7	3.9	40	2.9	4.3	15	1.1	1.5	3	0.2	0.2	9	0.6	0.6
-2	78	5.6	8.4	99	7.1	11.0	92	6.6	10.8	49	3.5	5.0	37	2.6	2.9	45	3.2	3.8
-1	177	12.6	21.0	155	11.0	22.0	168	12.0	22.8	146	10.4	15.4	175	12.5	15.3	165	11.8	15.6
0	279	19.9	40.9	262	18.7	40.7	236	16.8	39.6	231	16.5	31.9	369	26.3	41.6	439	31.3	46.9
+1	303	21.6	62.5	293	20.9	61.6	289	20.6	60.2	346	24.7	56.5	414	29.5	71.1	393	28.0	74.9
+2	288	20.5	83.0	266	19.0	80.5	267	19.0	79.3	311	22.2	78.7	310	22.1	93.2	247	17.6	92.5
+3	142	10.1	93.2	146	10.4	90.9	168	12.0	91.2	188	13.4	92.1	74	5.3	98.5	89	6.3	98.9
+4	68	4.8	98.0	87	6.2	97.1	73	5.2	96.4	80	5.7	97.8	19	1.4	99.9	16	1.1	100.0
+5	19	1.4	99.4	26	1.9	99.0	36	2.6	99.0	25	1.8	99.6	2	0.1	100.0			
+6	8	0.6	99.9	11	0.8	99.8	11	0.8	99.8	5	0.4	99.9						
+7	1	0.1	100.0	3	0.2	100.0	2	0.1	99.9									
+8							1	0.1	100.0	1	0.1	100.0						

NOTE: In some cases the cumulative percent (CUM%) of students shown on the table may not be the sum of the percent (%) column. These minor fluctuations are the results of rounding error.

This type of analysis was originally done for only the holistic scale; however, with the disturbing results which were revealed, it was determined that the same analysis should be done for the other five scales. This data is shown in the remainder of Table 3.

Locating the zero point (0) (Table 3) on the score change column and reading the cumulative percentage across that row we see that 40.9 percent of the students on the holistic scale, 40.7 percent for development, 39.6 percent for organization, 41.6 percent for sentence structure, and 46.9 percent of the students for mechanics showed no gain or a loss between pre/posttesting. Only the language development scale, with 31.9 percent at zero or less, was much better than the holistic scale, and even that figure might be considered too high a percentage of students to have shown no gain over a two year period of instruction.

This rather deflating news prompted speculation regarding the possible causes for the lack of growth, and the usual considerations were given to (1) reliability, (2) validity, (3) ceiling effect, and (4) floor effect. None of these possible explanations seem to have credence. Consequently, it was decided to analyze the results in order to determine which group of students were showing gains or no gains.

Table 4 displays the data relative to the 1,403 students with valid scores in another manner. The first column displays the change score scale (posttest minus pretest). Appearing across the top of the table are the 11 possible holistic pretest scores; and within the body of the table are the number of students with at a particular pretest score showing either a gain or loss relative to posttesting. Thus, for the one student whose change score was -6 we find that that individual had

a score of 7 on the pretest (and consequently scored a 1 on the posttest). The single student who recorded a change score of +7 (at the bottom of the table) had a pretest score of 3 (and obviously rose to a 10 on the posttest).

The data in Table 4 is divided into 4 quadrants with a fifth section representing those students who had no possibility of growth (all were at 10 on the pretest), and a sixth section of students who had no possibility of a score decline (all were at 0 on the pretest). The 20 students who had pretest scores of 0 on the pretest all improved. Four of the thirteen students, with pretest scores of ten, remained the same, and the remaining nine students regressed on the posttest at various degrees.

TABLE 4

Pretest/Posttest Score Comparisons
Senior Class of 1985 - N=1403
Holistic

Those Students Who Had Change Scores Of	Had Scores on the Prestest of										
	0	1	2	3	4	5	6	7	8	9	10
-6		**1**						1		**2**	**5**
-5						1		1	1		1
-4					1		5		4		1
-3						4	5	9	2	3	1
-2			2	3	8	10	22	10	15	6	2
-1		1	6	8	29	37	33	37	15	7	4
0		1	10	7	56	40	87	30	35	9	4
+1	1	5	8	28	69	51	78	35	18	10	
+2	4	4	20	27	88	32	64	27	22		
+3	5	4	15	22	26	25	25	20			
+4	6		12	5	19	9	17				
+5	2		4	1	6	6					
+6	2			1	5						
+7	**6**	**4**		1						**3**	

The 132 students in Quadrant 1 had scores of 4 or less on the pretest and all either stayed the same or regressed. These students are either not benefiting from the instruction they are receiving in writing skills, or they have not received instruction in writing skills to a degree sufficient to allow them to improve.

The 370 students in Quadrant 4 also had scores of 4 or less on the pretest; however, this group of students all increased on the posttest. The 429 students in Quadrant 2 all had relatively high scores on the pretest (from 5 to 9) and those scores remained constant or regressed, while those 439 students in Quadrant 3 had scores from 5 to 9 on the pretest and all increased their scores on the posttest.

COMMENTS

The first attempt to obtain pre/post writing samples of students from their beginning ninth grade to ending tenth grade years of schooling seems to have been relatively successful. While it would have been helpful to have obtained pre/post writing samples from a higher percentage of the total enrollment, the percent of students who did take the writing assessment (91.4 percent of the enrollment on the pretest and 85.9 percent on the posttest) and the number of valid scores obtained (80 percent of those taking the posttest had pretest scores which could be matched with their pretest scores) would seem to be satisfactory for the first assessment to be conducted by the district.

While the averages across all 6 score scales showed moderate overall gains for this group of students, further analysis of the data revealed that around 40 percent of the individual students failed to show any progress during their ninth and tenth grade years of schooling on 5 of the 6 score scales. On the remaining score scale approximately 32 percent of the students failed to show progress.

By cross-tabulating the change scores of students between their pre and posttests with their original pretest score, additional information was revealed:

1. Approximately 9 percent of the students had below 5 (0-10 scale) on the pretest and either remained at that level or regressed on the posttest.
2. Approximately 28 percent of the students had scores below 5 on on the pretest and improved their scores on the posttest.
3. Approximately 62 percent of the students scored from 5 to 9 on the pretest. When administered the posttest about half this group improved while the other half did not.

Findings 1 and 2 would seem to indicate that students entering ninth grade at the lower skill levels in composition are possibly given some elementary instruction in those skills since approximately 75 percent of that group seem to improve. However, fully 25 percent of that group either failed to improve or regressed by the end of the tenth grade.

Finding 3 shows a seemingly random nature in the balance between improvement and decline by the end of tenth grade among students who were average or above on the pretest. This would indicate that the teaching of composition skills to this group of students is either minimal or of such an elementary nature as to be ineffective. The higher order writing skills needed by this group of students are either not being taught or not being learned.

Most students of the future seem to agree that society is currently undergoing a major transformation that many equate with the magnitude of the transformation from an agriculturally oriented society to the industrially oriented society we have had since the mid-1800s. This new movement is being called, by most authorities, "The Information Age." The basis for this new "age" (according to the same authorities) is education and communication. There are only two standard means of communication dialogue: reading and writing, and listening and speaking. The acquisition of writing skills thus becomes one of the very roots of society and the economy when examined in this light. While communication skills have been (and are currently) extremely important, their importance will be magnified many times as we move into The Information Age.

With the increasing importance of writing skills within our society, it would seem that a continuing program designed to monitor student acquisition of those skills would be a logical need. Experience has shown that the acquisition of skills is a much more difficult and time-consuming activity than the acquisition of knowledge; therefore, an annual assessment of writing skills would probably be more frequent than necessary.

It is the recommendation of the Department of Evaluation and Research that the Instructional Division consider the need for and timing of student composition assessments, and make recommendations to this department for their conduct, in sufficient time for such tests to be designed, validated, and placed in use when desired.

APPENDIX A

Writing Assessment Booklet

WRITING ASSESSMENT BOOKLET

DEPARTMENT OF EVALUATION AND RESEARCH
DES MOINES PUBLIC SCHOOLS
DES MOINES, IOWA

STUDENT NAME: (please print) (1) LAST (14) (15) FIRST (20) GRADE: (21)

STUDENT ID NUMBER: (23) (28) SCHOOL NUMBER: (29-31)

TEACHER LAST NAME: (32) (45) DATE: _____

	FIRST READER ____	SECOND READER ____
HOLISTIC SCORE	_____ (46)	_____ (47)
COMPOSITION TRAITS		
1. Development Details are relevant Details are specific	5 4 3 2 1 0 (48)	5 4 3 2 1 0 (49)
2. Organization Ideas are ordered Ideas are related to the assigned topic Transitional devices are used	5 4 3 2 1 0 (50)	5 4 3 2 1 0 (51)
3. Language Choice Appropriate word choice Variety of word choice Richness of expression	5 4 3 2 1 0 (52)	5 4 3 2 1 0 (53)
4. Sentence Structure Sentences are varied Sentences express complete thought	3 2 1 0 (54)	3 2 1 0 (55)
5. Mechanics Spelling Punctuation Usage Capitalization Agreement	3 2 1 0 (56)	3 2 1 0 (57)

STUDENT INSTRUCTIONS

In this writing assessment you will be asked to write a legible, well-organized, grammatically correct essay on a specified task. Please read these instructions thoroughly before you begin writing.

You will be given scratch paper to plan your composition and to write your first draft. Write your final draft in this booklet in the space provided. Please limit the length of your composition to the space available in the lined pages of this booklet.

When you write your essay be sure to keep the following points in mind:

1. Give facts that clearly support your position.
2. Write your ideas in logical order.
3. Choose words that say exactly what you mean.
4. Proofread your essay carefully for spelling, punctuation, capitalization and usage errors.
5. Allow yourself time for recopying the essay in your booklet.
6. You shall have a total of 45 minutes for planning, writing, and recopying your essay. Your teacher will warn you when you have only fifteen minutes remaining.

Carefully read the writing assignment in the box below; then begin.

> Assume some legislators are discussing raising to 18 the age for earning a driver's license. Some reasons cited are the gas shortage, high accident rates, cost of drivers education classes, and poor academic performance in schools.
>
> React to this problem by writing an essay agreeing or disagreeing with the proposal to raise the driving age in Iowa to 18.

APPENDIX B

Holistic Scoring Traits
Composition Scoring Traits

HOLISTIC SCORING TRAITS

Holistic Score	
10	A superior response will not just take a position on the issue but will also list reasons and support for those reasons. It will explain why each reason supports the position. A superior paper will be literate and orderly.
8	This score will be useful for a well-handled paper which is deficient in one or two characteristics of the superior response, ie. it may emphasize only one or two reasons, and the paper may be less fully developed. The paper shows overall competence in writing.
6	This score will indicate a paper which addresses the topic but may not take a definite stand (or discusses both sides of the topic). The reason or reasons mentioned are not as well developed as #4. Mechanical errors are obvious and repeated.
4	In attempting to take a stand, the writer of the paper assumes the reader agrees with him, and, therefore, presents little explanation of his position or why he believes what he does. Mechanical problems are repeated and distract from any point the paper may make. This writer appears to have started in the middle of the task.
2	This paper presents a topic related to, but not the same as, the task suggests. The treatment is superficial or stereotyped. The writing exhibits several weaknesses in spelling, punctuation wording.
0	This score is to be used for papers which are lacking in focus and substance, depart completely from the assigned topic, and/or exhibit serious writing faults.
*	Non-response papers and papers which do not fall into the foregoing categories, extensively agree with the question, or are otherwise idiosyncratic, should immediately be brought to the attention of the table reader and the question leader.

COMPOSITION SCORING TRAITS

	Development	Organization	Language Choice
10	All supporting details are specific, consistent, relevant. All generalizations are developed with specific details.	Ideas are ordered and progress logically. Ideas are related to assigned topic. Transitional devices are used.	Sophisticated expression, use of synonyms rather than repetition, diction, comparisons used effectively.
8	Most details are specific and relevant.	Ideas are ordered and progress logically, most ideas are related to assigned topic. Little or no transition.	Use of varied expressions, synonyms rather than repetition, expanded vocabulary.
6	Mostly generalizations used. Few specific details used. Details are relevant.	Progression of ideas. Paragraphs used. Weak sense of beginning, middle and end. Lacking transitional devices.	Attempt to use expanded vocabulary. Some use of adjectives, adverbs, action oriented verbs.
4	Lack of progression of ideas. Details not relevant. Details not specific.	Lack of paragraphing but some sense of beginning, middle and end. Lacking transitional devices.	Generally lack attempt to go beyond. Common words repetition, no use of synonyms.
2	Used only generalizations, no specific examples, details not relevant.	7 or 8 sentences but no order, no indention, no sense of beginning middle or end. Lacking transitional devices.	Limited vocabulary.
0	Has no development of details.	Lack of paragraphing. No beginning, middle or end. Very short - one or two sentences.	Very limited vocabulary. Words misused.

	Sentence Structure	Mechanics Traits
6	Complete sentences used. Good variety of sentence patterns used.	Generally correct spelling; uses a variety of punctuation well; verb tense consistent; pronouns usually agree in number; subjects and verbs agree; apostrophes indicate possession generally neat and written in paragraphs.
4	Complete sentences used for the most part but very little variety in sentence patterns. Some awareness of sentence structure but some fragments and run-ons occur.	Misspelling of common words and phonetical attempts spell difficult words; comma splices and run-ons with fragments; inconsistent end punctuation; pronoun shifts; apostrophe use inconsistent. Some crossing out and written in short paragraphs or with no paragraph division.
2	Some sentence sense, but fragments and run-ons occur.	Misspells by omitting vowels; leaves out words; sentences are often fragments or connected by commas; capitals misused; tense shifts; number shifts; apostrophe misused. Generally messy and likely be short.
0	Very little sentence sense - many fragments, run-on sentences.	Misspelling indicates a lack of symbol/ sound correspondence; omitted words; misuse of capitals; tense shifts; number shifts; apostrophe misused. Generally messy and likely to be short. May be incomprehensible.

APPENDIX C

Prototype Papers

I agree with the proposal to raise the legal driving age in Iowa to 18. I feel that raising the driving age to 18 is a wise choice because it will cut down on the use of cars, the high accident rates, the cost of driver education classes and the poor academic performances in schools.

By cutting down on the use of cars, we will be cutting down the gas consumption of Iowans. Because of the fact that there will be excess gas the prices may be lowered. Many students will probably turn to bus lines, increasing the city's revenue.

Teenagers on drugs or who drink won't be able to drive because of their age which means that high accidents rates will probably drop.

Changing the driving age to 18 will also cut down the need for driver education classes in the schools since you must be 18 years old to have a drivers license. The money that is used now for driver education classes can be put to better uses.

Without cars many students won't be able to hang out with friends or scoop the loop. The students will be at home, so they will have more time to study. This should raise the academic performance in schools therefore producing higher educated people who can have a better chance at jobs in the future.

Many people may complain about raising

the driving age, but as you can see it will be very worthwhile, because we will be solving four problems with one small change. the gas shortage, the high accident rates, the cost of driver education classes, and poor academic performances.

I disagree with the legislators on the subject of raising the age for earning a driver's license to 18, for many reasons.

My main argument is that their reasoning is faulty. They state one of their reasons is poor academic performance in schools. I would like to know how they got the evidence that puts driving as the cause. I say that there is no way to tell that driving should be blamed because it depends upon what kind of person is performing poorly, there are too many different reasons for why they do so.

A gas shortage is another one of their reasons. This one is not hard to put down at all. Many 16 year-olds car-pool to school which saves gas. I thought that's what the government wants us to do.

High accident rates is their only reason that makes any sense, but it is not a good basis on which to support the proposal. I too believe that accidents are a big problem but there is a better way to slow down

these rates by insisting that there be better testing for persons applying for a licenses. There should be classes offered on what to do in accidental situations and how to avoid them.

In closing, I would like to also add that if we insist on stricter rules and guidelines for driving, there will be no need to raise the age for earning a license to 18.

6

Raising the driving age to eighteen will stop the problem of gas shortage, but will put put kids under eighteen out of jobs. And start other problems.

If a person under the age of eighteen wants to drive and is willing to pay the cost of drivers education class they should be allowed that privilage.

Poor academic performances in school shouldn't be the case of worry. Poor driving should be.

If a student can't get drivers education he (or she) might drive without his or her license. Which could be more dangerous to everyone else including the driver without the license.

When a person under 18 is allowed the right to drive and then they make a new law. What do they do with the kid whose sixteen and already has a

license and a car? He (or she) has spent more money on insurance and the needs of the car than it would cost them for drivers education. Then they make a law breaker.

Who are they really trying to help? ~~the~~ It sure isn't the 16 year old!

NO I dont think that the age should be raised to 18, because at 16 most teenagers ought to be responsible, and able to keep their grades up and things. And it wouldnt be fair to other kids who are ready to take that responsibility to have a car and not get high and drunk and stuff. I feel that some teenagers might have a chance to get a car and we wont be able to drive it because of the age limit on driving We should have the rights to go places, not walking & others but driving. And not having to depend upon our mothers, fathers and freinds & older than us to take us places, I think that is also learning responsability, And thats why that I think that they shouldnt change the age limit to 18.

I disagree for the legislators to raise to get your driver's licence at 18. The only reason they are doing this is to save money. What does education has to do with it. Stat problem how we do in drive. There has never been a gas shortage in Iowa by far. They do not like the way we drive. Overall. I think we drive better than old people. all the accident you see or hear drive's from age 20 and on. It as been a problem why didn't you do this before now. why do we take drive education for to learn how to drive. and how do we past the driver test.

I do not think the driving law should not be raised because people who had jobs. Could not get to work and it would not be fair because the gas shortage. because people would use all the gas up before we get our license and if you just got your license and you were sixteen and a car and told you could not drive it no more, would't you be mad I do not agree.

A P P E N D I X D

Comparison of Enrollments and Number of Students Taking the Writing Assessment
Fall 1981 - Spring 1983

TABLE 5

Ninth Grade Fall, 1981, Enrollments
Compared to Number of Students Taking the Writing Assessment

Building	Fall, 1981 Enrollment	# Students Taking Composition Test	% of Students Taking Test
East High School	493	462	93.7
Hoover High School	339	319	94.1
North High School	206	213	103.4
Roosevelt High School	301	298	99.0
Technical High School	437	348	79.6
Brody Junior High School	182	173	95.1
Douglass Learning Center	29	29	100.0
McCombs Junior High School	197	171	86.8
Weeks Junior High School	221	185	83.7
Total	2405	2198	91.4

TABLE 6

Tenth Grade Spring, 1983, Enrollments in English
Compared to Number of Students Taking the Writing Assessment

Building	Fall, 1981 Enrollment	# Students Taking Composition Test	% of Students Taking Test
East High School	393	304	77.4
Hoover High School	323	303	93.8
Lincoln High School	513	474	92.4
North High School	169	125	74.0
Roosevelt High School	296	261	88.2
Technical High School	354	293	82.8
Total	2048	1760	85.9